UPAR7: A knowledge-based system for headline sentiment tagging

François-Régis Chaumartin
Lattice/Talana – Université Paris 7
30, rue du château des rentiers - 75013 Paris - France
fchaumartin@linguist.jussieu.fr / frc@proxem.com

Abstract

For the Affective Text task at SemEval-2007, University Paris 7's system first evaluates emotion and valence on all words of a news headline (using enriched versions of SentiWordNet and a subset of WordNet-Affect). We use a parser to find the head word, considering that it has a major importance. We also detect contrasts (between positive and negative words) that shift valence. Our knowledge-based system achieves high accuracy on emotion and valence annotation. These results show that working with linguistic techniques and a broad-coverage lexicon is a viable approach to sentiment analysis of headlines.

1 Introduction

1.1 Objectives

The detection of emotional connotations in texts is a recent task in computational linguistics. Its economic stakes are promising; for example, a company could detect, by analyzing the blogosphere, people's opinion on its products.

The goal of the SemEval task is to annotate news headlines for emotions (using a predefined list: anger, disgust, fear, joy, sadness & surprise), and for valence (positive or negative). A specific difficulty here is related to the small number of words available for the analysis.

1.2 Overall architecture

Our system is mainly rule-based and uses a linguistic approach. From a macroscopic point of view, we follow the hypothesis that, in a news title, all the words potentially carry emotions. If linguistic resources make it possible to detect these emotions individually, how can we deal with headlines where bad and good emotions appear at once?

Our objective is to identify the expression which carries the main topic of the title. One can consider that this expression has a primary importance.

We also seek to lay down rules for detecting specific emotions. For instance, surprise sometimes comes from the contrast between good and bad news. And sometimes, simple lexical elements are characteristic of an emotion; a negation or a modal auxiliary in a title may be a relevant indicator of surprise.

We describe here the techniques we implemented to address all these points.

2 Components & resources used

The system we employed for the Affective Text evaluation consists of the following components[1]:

- The SS-Tagger (a Part-of-Speech tagger)[2],
- The Stanford Parser.

We also used several lexical resources:

- WordNet version 2.1,
- A subset of WordNet-Affect,
- SentiWordNet.

As the SS-Tagger is straightforward, we will not say more about it here. We will, however, discuss the remaining components and resources below.

[1] We used them through the Antelope NLP framework (www.proxem.com), which makes them easy to use.
[2] This fast PoS tagger uses an extension of Maximum Entropy Markov Models. See (Tsuruoka, Tsujii, 2005).

Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007), pages 422–425,
Prague, June 2007. ©2007 Association for Computational Linguistics

2.1 Choice of the Stanford Parser

We wished to use a syntactic parser for this task. We hesitated between two parsers producing a dependency graph, the *Link Grammar Parser* (Sleator, Temperley, 1991) and the *Stanford Parser* (Manning, Klein, 2002).

As a news title is sometimes reduced to a nominal group, without a verb, our experiments showed that we should modify the title to make it "grammatically correct". Such a step is essential to obtain accurate results with a rule-based analyzer such as the Link Grammar Parser. On the other hand, a statistical analyzer like the Stanford Parser is more tolerant with constructions which are not grammatically correct. That is why we chose it.

2.2 WordNet

We used WordNet (Miller, 1995) as a semantic lexicon. This well-known project, started in 1985 at Princeton, offers a broad-coverage semantic network of the English language, and is probably one of the most popular NLP resources.

In WordNet, words are grouped into sets of synonyms. Various semantic relations exist between these synsets (for example, hypernymy and hyponymy, antonymy, derivation…).

2.3 WordNet-Affect

WordNet-Affect (Strapparava, Valitutti, 2004) is a hierarchy of "affective domain labels", with which the synsets representing affective concepts are further annotated. We used the subset of WordNet-Affect provided as emotions lists by the SemEval organizers. To improve it, we manually added to the emotion lists new words that we found important on the task trial data.

	Nouns	Verbs	Adjectives	Adverbs
Anger	37	26	16	0
Disgust	35	19	9	0
Fear	71	26	20	4
Joy	50	22	14	1
Sadness	88	37	29	4
Surprise	16	29	13	2

Table 1: Counting of new words for each emotion.

Nouns	Verbs	Adjectives	Adverbs
cancer	demolish	comatose	bloody
danger	injure	nuclear	dead
poverty	kidnap	violent	worse

Table 2: Some words added for "fear" emotion.

The synsets of emotions lists were considered as seeds; our system recursively propagated their emotions to their neighbor synsets[3].

SentiWordNet

SentiWordNet (Esuli, Sebastiani, 2006) describes itself as a lexical resource for opinion mining. SentiWordNet assigns to each synset of WordNet three sentiment scores[4]: positivity, negativity, objectivity, the sum of which always equals 1.0.

This resource has been created with a mix of linguistics and statistics (using classifiers). The advantage of this approach is to allow the automatic generation of emotion values for all the synsets of WordNet. The disadvantage is that, as all the results are not manually validated, some resulting classifications can appear incorrect[5].

We recursively propagate the positivity and negativity values throughout neighbor synsets[6].

3 UPAR7 Affective Text system

3.1 "De-capitalization" of common words

A preliminary problem that we had to solve was related to the Anglo-Saxon habit of putting initial capital letters in all the words of a title.

The first pass of our system thus detected news titles that were "improperly" capitalized, and "decapitalizes" their common words.

For that, we used the SS-Tagger on the title; according to the part of speech of each word, information found in WordNet, and some hand-crafted rules[7], the system chooses or not to keep the initial.

The impact of this processing step is far from negligible, from the point of view of the Stanford Parser. Indeed, let us have a look at the difference between the parsing of the title, before (figure 1) and after (figure 2) this processing.

[3] Following relations such as Hyponym, Derivation, Adjective Similar, Adjective Participle, Derivation and Pertainym.

[4] For instance, the synset ESTIMABLE#1 (*deserving of respect or high regard*) has: Positivity = 0.75, Negativity = 0.00, Objectivity = 0.25.

[5] For example, RAPE#3 (*the crime of forcing a woman to submit to sexual intercourse against her will*) is classified with Positivity=0.25 and Negativity=0.0 despite the presence of the word "crime" in its gloss.

[6] Using WordNet's relations such as Hyponym (for noun and verb), Antonym and Derivation. For antonyms, positivity and negativity values are exchanged.

[7] For instance, a word that cannot be any form of a WordNet lemma is probably a proper noun, and then we keep its initial.

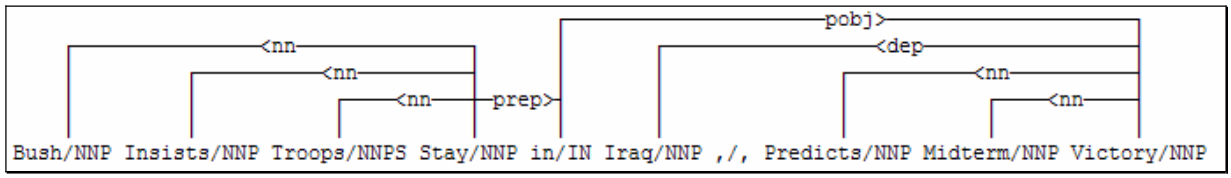

Figure 1 : Output of the Stanford Parser with a title that is "**improperly"** capitalized.

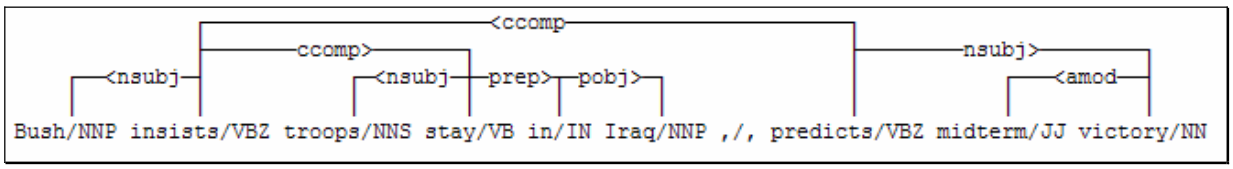

Figure 2 : Output of the Stanford Parser with a title that is "**properly"** capitalized.
(Words are tagged with the right part-of-speech, and dependencies are now correct.)

3.2 Individual words rating

For the moment, we consider the output of the Stanford Parser as an array of PoS-tagged words. We use WordNet's morphology functions to find the possible base form of each word.

At this stage, an important question arose: was lexical disambiguation possible? We thought not, because with short sentences, few relevant heuristics apply. We chose another solution, by considering that the emotion and valence values of a word were the linear combination of that of all its possible meanings, balanced by the frequency of each lemma.

We detected emotion and valence values for each word, by using our enriched version of WordNet-Affect and SentiWordNet.

In fact, we also detected some extra information:

- An additional 7th emotion, that looks like "compassion for people needing protection". Our assumption is that certain words express a subjacent need for protection. For example, there is "student" behind "school", and "child" behind "adoption". So, we built a list of words designating something that needs protection; we also include in this list words such as "troops", "shoppers"…
- We tried to detect acronyms relating to technology; for this, we defined a list of high-tech companies and a very basic regular expression rule saying that a word (not in WordNet) containing numbers, or capitals not in first position, should be something high-tech. (This very basic rule seems to work nicely on PS3, iPod, NASA…). We use these high-tech indications to increase the "joy" emotion.
- We counted lexical elements that we think are good indicators of surprise: negations, modal auxiliaries, question marks.

At this stage, we begin some post-processing on individual words. Which factors cause anger rather that sadness? We believe that human intention (to harm) causes the former emotion, while natural factors such as disease or climatic catastrophes cause the latter. So, we used a few rules related to the WordNet noun hierarchy, based on the fact that when a noun is a hyponym of a given synset, we boost some emotions:

Does noun inherit from?	Emotions to boost
UNHEALTHINESS	Fear, sadness
ATMOSPHERIC PHENOMENON	Fear, sadness
AGGRESSION, HOSTILITY, WRONGFUL CONDUCT	Anger, fear, sadness, disgust
WEAPONRY, WEAPON SYSTEM	Anger, fear, sadness
UNFORTUNATE PERSON	Sadness, "compassion"
HUMAN WILL	Anger

Table 3: Hypernyms triggering an emotion boost.

Then, the emotions found serve to update the valence, by increasing positivity or negativity:

Emotion	Positivity	Negativity
Joy	++	--
Anger, disgust, sadness, fear, "compassion"	--	++

Table 4: Emotions that change valence.

3.3 Global sentence rating

At this stage, our system tries to find the main subject of the news title. Again, we use the output of the Stanford Parser, but this time, we make use of the dependency graph. We consider that the main word is the root of the dependency graph, i.e. the word that is never a dependant word. (For instance, in figure 2, the main word is "predicts".)

We think that the contribution of this main word is much more important than that of the other words of the title[8]. So, we multiply its individual valence and emotion by 6.

The last important part of linguistic processing is the detection of contrasts and accentuations between "good" or "bad" things. We search patterns like [noun→subject→verb] or [verb→direct object→noun] in the dependency graph, with verbs that increase or decrease a quantity[9]. Using the valence of the given noun, this gives our system the ability to detect very good news ("boosts (brain) power") or good news where something bad gets less important ("reduces risk", "slows decline", "hurricane weakens"…).

4 Results

	Fine-grained	Coase-grained		
	Pearson	Accuracy	Precision	Recall
Anger	32.33	93.60	16.67	1.66
Disgust	12.85	95.30	0.00	0.00
Fear	44.92	87.90	33.33	2.54
Joy	22.49	82.20	54.54	6.66
Sadness	40.98	89.00	48.97	22.02
Surprise	16.71	88.60	12.12	1.25

Table 5: Results of the emotion annotation.

Our rule-based system detects the six emotions in news headlines with an average accuracy reaching 89.43% (coarse-grained evaluation). However, recall is low.

	Fine-grained	Coase-grained		
	Pearson	Accuracy	Precision	Recall
Valence	36.96	55.00	57.54	8.78

Table 6: Results of the valence annotation.

The valence detection accuracy (55% in coarse-grained evaluation) is lower than in emotion annotation. We attribute this difference to the fact that it is easier to detect emotions (that are given by individual words) rather than valence, which needs a global understanding of the sentence.

5 Conclusion

Emotion and valence tagging is a complex and interesting task. For our first attempt, we designed and developed a linguistic rule-based system, using WordNet, SentiWordNet and WordNet-Affect lexical resources, that delivers high accuracy results. In our future work, we will explore the potential of simultaneously using a statistical approach, in order to improve recall of sentiment annotation.

References

Andrea Esuli, Fabrizio Sebastiani. 2006. *SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining.* Proceedings of LREC 2006, fifth international conference on Language Resources and Evaluation, pp. 417-422.

Christopher Manning, Dan Klein. 2002. *Fast Exact Inference with a Factored Model for Natural Language Parsing*. Advances in Neural Information Processing Systems 15 (NIPS 2002).

George Miller. 1995. *WordNet: A lexical database*. Acts of ACM 38, 39-41.

Livia Polanyi, Annie Zaenen. 2006. Contextual Valence Shifters. In J. G. Shanahan, Y. Qu, and J. Wiebe, editors, *Computing Attitude and Affect in Text: Theory and Application*. Springer Verlag.

Daniel Sleator, Davy Temperley. 1991. *Parsing English with a Link Grammar*. Acts of Third International Workshop on Parsing Technologies.

Carlo Strapparava, Alessandro Valitutti. 2004. *WordNet-Affect: an affective extension of WordNet.* Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004), pp. 1083-1086.

Yoshimasa Tsuruoka, Jun'ichi Tsujii. 2005. *Bidirectional Inference with the Easiest-First Strategy for Tagging Sequence Data*. Proceedings of HLT/EMNLP 2005, pp. 467-474.

[8] In sentences like "study says…", "scientists say…", "police affirm…", the main head word is the verb of the relative.

[9] We "rediscovered" valence shifters (words that modify the sentiment expressed by a sentiment-bearing word, see (Polanyi and Zaenen, 2006)).